***Journal of Physical Activity and Health*, (Ahead of Print)**
https://doi.org/10.1123/jpah.2017-0089

Human Kinetics
ORIGINAL RESEARCH

Understanding the Nature of Measurement Error When Estimating Energy Expenditure and Physical Activity via Physical Activity Recall

David R. Paul, Ryan McGrath, Chantal A. Vella, Matthew Kramer, David J. Baer, and Alanna J. Moshfegh

Background: The National Health and Nutrition Examination Survey physical activity questionnaire (PAQ) is used to estimate activity energy expenditure (AEE) and moderate to vigorous physical activity (MVPA). Bias and variance in estimates of AEE and MVPA from the PAQ have not been described, nor the impact of measurement error when utilizing the PAQ to predict biomarkers and categorize individuals. **Methods**: The PAQ was administered to 385 adults to estimate AEE (AEE:PAQ) and MVPA (MVPA:PAQ), while simultaneously measuring AEE with doubly labeled water (DLW; AEE:DLW) and MVPA with an accelerometer (MVPA:A). **Results**: Although AEE:PAQ [3.4 (2.2) MJ·d^{-1}] was not significantly different from AEE:DLW [3.6 (1.6) MJ·d^{-1}; $P > .14$], MVPA:PAQ [36.2 (24.4) min·d^{-1}] was significantly higher than MVPA:A [8.0 (10.4) min·d^{-1}; $P < .0001$]. AEE:PAQ regressed on AEE:DLW and MVPA:PAQ regressed on MVPA:A yielded not only significant positive relationships but also large residual variances. The relationships between AEE and MVPA, and 10 of the 12 biomarkers were underestimated by the PAQ. When compared with accelerometers, the PAQ overestimated the number of participants who met the Physical Activity Guidelines for Americans. **Conclusions**: Group-level bias in AEE:PAQ was small, but large for MVPA:PAQ. Poor within-participant estimates of AEE:PAQ and MVPA:PAQ lead to attenuated relationships with biomarkers and misclassifications of participants who met or who did not meet the Physical Activity Guidelines for Americans.

Keywords: accelerometry, epidemiology, physical activity assessment

The significance of the relationship between physical activity and health has been well documented[1–4] and has led to the development of public health efforts to increase physical activity, moderate to vigorous physical activity (MVPA) in particular.[5] Similarly, considerable efforts have been devoted to population-level physical activity surveillance, such as the National Health and Nutrition Examination Survey (NHANES),[6] which provides a comprehensive understanding of the status of physical activity levels (PALs) in the United States and the subsequent relationships between physical activity and clinically relevant health outcomes. Reasonably, accurate estimates of physical activity from surveillance studies are critical given their role in developing physical activity guidelines, such as the Physical Activity Guidelines for Americans (PAGA).[5]

Despite the importance of conducting both physical activity surveillance and promotion studies, investigators continue to struggle with the complexities associated with physical activity measurement and assessment.[7–9] Evidence demonstrating the associations between energy expenditure, physical activity, and health has been derived from both objective and subjective techniques.[7,9,10] Two of the more widely adopted objective techniques are doubly labeled water (DLW), which provides an estimate of total energy expenditure (TEE) for a 7- to 14-day period,[11,12] and accelerometers, which provide a detailed picture of free-living physical activity patterns.[13] Although DLW and accelerometers have been used extensively, these techniques pose some disadvantages, such as cost, participant burden, within-participant measurement error, and short time frame of measurement.[10,12,14,15]

Physical activity questionnaires (PAQs) have long been used to estimate energy expenditure and physical activity in a manner that is simple to administer, low cost, and imposes minimal participant burden.[16] Unfortunately, PAQs have been shown to both overestimate and underestimate energy expenditure and physical activity.[9,14,17–20] These inaccuracies could have implications for how the data are applied in subsequent statistical analyses, such as correctly characterizing the number of individuals who meet the PAGA.[5]

Despite these limitations, much of what is known about the connection between physical activity and health is derived from techniques, such as PAQs.[7] An important question is whether "subjective" techniques that may lack precision (such as PAQs) can be used to establish true associations between energy expenditure, physical activity, and disease. Although several PAQs have undergone a validation process,[21–24] these validation studies also have limitations.[25] For example, some validation studies have implemented incorrect statistical approaches, made comparisons using subjective measures as the criterion, had small sample sizes, and/or failed to report within-participant measurement error.[9,14,21,23,25] Concerns have also been raised about whether the error observed when comparing a subjective technique to an objective one can solely be attributed to the subjective technique.[26] Finally, although predictions of energy expenditure and physical activity from subjective techniques have often been compared with

Paul and Vella are with the Dept of Movement Sciences, University of Idaho, Moscow, ID. McGrath is with the Dept of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI. Kramer is with the Biometrical Consulting Service, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD. Baer is with the Food Components and Health Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD. Moshfegh is with the Food Surveys Research Group, Beltsville Human Nutrition Research Center, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD. Paul (dpaul@uidaho.edu) is corresponding author.

more accepted objective techniques, few studies have assessed the impact of using subjective measures to predict biomarkers of disease.[27] For example, despite the presence of considerable measurement error, it is possible that PAQs may still adequately predict disease risk by "ranking" participants accurately.[7,28]

The purpose of this study was to understand how self-reported energy expenditure and physical activity may be overestimated or underestimated by comparing estimates of AEE between PAQ and DLW, and MVPA between PAQ and accelerometer. Furthermore, we take the results a step further by quantifying the implications of applying reported energy expenditure and physical activity results, such as understanding the relationships with biomarkers and characterizing individuals who meet the PAGA.

Methods

Participants

Prior to study enrollment, participants completed a screening visit that included a medical history; cardiometabolic profile (complete blood count and analysis of fasting blood and urine samples); and measurement of height, weight, and blood pressure. All participants were weight stable; not actively pursuing weight loss; and not taking medications known to affect food intake, appetite, or water balance. Those excluded from the study were pregnant and lactating females, persons with diabetes, and nutrition professionals. After screening, a total of 262 women and 262 men agreed to participate in the study and provided written informed consent. The study protocol was approved by the Johns Hopkins Bloomberg School of Public Health Committee on Human Research.

Study Design

Participants completed 4 visits during a 2-week period at the Beltsville Human Nutrition Research Center, as described previously.[29] At the first visit, participants provided a urine sample, drank a dose of DLW for the measurement of TEE, and were given an accelerometer. Weight was measured to the nearest 0.01 kg, and height was measured to the nearest 0.10 cm. Body mass index was calculated using body weight in kilograms divided by height in meters squared. Body composition was measured by dual-energy X-ray absorptiometry (QDR 4500, Hologic Inc, Marlborough, MA).[30] The second and third visits were scheduled 5–6 and 10–11 days, respectively, after the first visit. Spot urine samples for the determination of TEE were collected daily by participants during the second week of the study. At visits 2 and 3, participants provided spot urine samples and returned any urine samples collected at home. On the 10th or 11th day of the study, the PAQ was administered by a trained interviewer (Westat Corp, Rockville, MA) over the telephone. At visit 4 (approximately 14 d after the first visit), participants provided a final urine specimen, were measured for resting energy expenditure (REE) via indirect calorimetry,[29] and returned the accelerometer.

Measurement of TEE and MVPA via PAQ

The NHANES PAQ (2001–2002 version) queries participants about their physical activities from the previous 30 days, including walking and biking to and from work or school; tasks performed in or around the home or yard that required moderate effort; usual daily activities; physical activities (eg, exercise, sports, and physically active hobbies); and activities designed to strengthen muscles. Activities performed for at least 10 minutes were categorized as "moderate" or "vigorous," then summed to derive reported minutes of MVPA via PAQ (MVPA:PAQ). The values for MVPA were then used to determine if each participant met the PAGA of at least 150 minutes per week of MVPA in bouts ≥10 minutes.[5,6,17] TEE was calculated by multiplying a corresponding metabolic equivalent (MET) code (1 MET = 4.184 $kJ \cdot kg^{-1} \cdot h^{-1}$) by the time spent in each activity over the previous 30 days. In cases where respondents reported no physical activity, TEE was calculated by multiplying REE by a PAL (TEE/REE) of 1.69 for women and 1.64 for men.[31] All values were standardized to a per day basis, assuming a month has 30 days.

Preliminary results indicated that some of the TEE estimates were implausible (eg, >80 $MJ \cdot d^{-1}$). Therefore, estimates of upper and lower bounds for plausible TEE values were developed based on PAL as a means to exclude participants from further analysis. First, the plausible range of PAL for weight-stable individuals was estimated to be 1.2–2.2; this range was derived from the midpoint of the lowest (sedentary) and highest (very active) PAL ranges defined by the Institute of Medicine.[32] Then, based on replicate measures of TEE (via DLW) and REE in 40 weight-stable adults from our laboratory, total measurement error in PAL was estimated to be 5.947%. The estimate for measurement error in PAL and a sample average PAL value of 1.691 were used to compute a 95% confidence interval width of 0.197. Finally, the 0.197 confidence interval width was applied to the 1.2–2.2 plausible PAL range to obtain a lower PAL cutoff of 1.003 and an upper PAL cutoff of 2.397.

Measurement of TEE via DLW

On visit 1, participants drank a previously mixed dose of DLW containing approximately 0.10 g of H_2O per kilogram and 0.08 g of $H_2{}^{18}O$ per kilogram of body weight.[29] A 24-hour urine sample collected on the previous day was used to measure background isotope enrichments. Isotopic enrichment of urine samples was measured using continuous-flow isotope-ratio mass spectroscopy (Europa Scientific Hydra, Crewe, United Kingdom). The spectrometer was calibrated before the analysis of each participant's sample, which was analyzed in triplicate.

Measurement of MVPA via Accelerometer

An ActiGraph 7164 accelerometer (ActiGraph Corp, Pensacola, FL) was worn on a snuggly fitting waist belt that was set to store the data in 60-second intervals. Participants were asked to wear the accelerometer on their right hip (unless they reported being unable), so each individual consistently wore it on the same side and location. Participants were also asked to maintain an activity log that recorded their sleeping time, waking time, accelerometer removal, and any engagement in structured exercise. Participants who failed to wear the accelerometer for a minimum of 7 days for at least 12 hours per day were excluded from the analysis.[33] Time spent in MVPA via accelerometer (MVPA:A) was defined as the number of minutes spent ≥2020 counts per minute in bouts ≥10 minutes (with allowances for 2 consecutive minutes below the threshold) and subsequently used to determine if each participant met the PAGA.[5,6,17]

Analysis of Blood Samples

Fasting blood samples were collected at visit 1 and analyzed for high-density lipoprotein and triglycerides. All samples were sent to the MedStar's Penn Medical Laboratory (Washington, DC) for analysis. These lipids were chosen based on their well-established associations with aerobic exercise.[34]

Statistical Analyses

Activity energy expenditure (AEE) was calculated as 0.9 × (TEE − REE) for both the PAQ (AEE:PAQ) and DLW (AEE:DLW).[28] Paired *t* tests were conducted to compare differences between AEE:PAQ and AEE:DLW, and MVPA:PAQ and MVPA:A. Regression analyses were conducted to estimate the relationships between AEE:PAQ and AEE:DLW, and MVPA:PAQ and MVPA:A. Additional explanatory variables were explored to determine their influence on the relationships between AEE:PAQ and AEE:DLW, and MVPA:PAQ and MVPA:A, based on their ability to reduce the residual variance and Akaike's information criteria. The presence of attenuation (underestimation) in the relationships between AEE or MVPA and a variety of biomarkers was detected by calculating correlation coefficients and *P*-values of PAQ results when compared with those from DLW and accelerometers (serving as the criterion standards). Correlation coefficients were calculated with log-transformed data to better meet linear models assumptions, such as homogeneity of variance. Results are presented as the mean (SD). Statistical significance was set at an alpha level of .05 for all analyses. SAS software (version 9.4; Cary, NC) was used for all the analyses.

Results

Of the 524 participants who provided consent, 27 were excluded for noncompliance or inconsistent DLW data, and 11 were excluded for not meeting the minimum accelerometer wear time criteria. After exclusions due to implausible PAQ data (n = 101), the results from 385 participants were included in the analyses (Table 1). The participants excluded due to implausible results were 50.7 (11.8) years of age, 30.6% (7.6%) of body fat, 27.3 (4.3) kg·m^{-2} of body mass index, and engaged in 14.4 (21.9) minutes per day of MVPA and 4.4 (1.8) MJ per day of AEE.

Table 1 Descriptive Characteristics of the Participants (n = 385) Included in the Analyses

Sex, women/men	207/178
Age, y	49.0 (10.6)
Body fat, %	32.6 (7.6)
Body mass index, kg·m^{-2}	26.3 (4.5)
Fat mass, kg	25.0 (8.7)
Fat-free mass, kg	51.0 (10.9)
High-density lipoprotein, mg·dL^{-1}	50.0 (14.4)
Triglyceride, mg·dL^{-1}	98.0 (55.7)

Note. Data are presented as mean (SD).

The average AEE:PAQ and AEE:DLW were 3.4 (2.2) MJ per day and 3.6 (1.6) MJ per day, respectively ($P > .14$), with an average absolute difference of 2.1 (1.5) MJ per day. The average MVPA:PAQ and MVPA:A were 36.2 (24.4) minutes per day and 8.0 (10.4) minutes per day, respectively ($P < .0001$), with an average absolute difference of 28.7 (22.0) minutes per day. Absolute differences for AEE and MVPA are shown in Figure 1. Scatter plots of AEE:PAQ and AEE:DLW, and MVPA:PAQ and MVPA:A are shown in Figure 2. The relationship between AEE:PAQ and AEE:DLW was positive ($b_1 = 0.15$; megajoules per day) and significant ($P = .04$) with a residual variance of 4.9. After adjustment for fat-free mass ($P < .0001$), the relationship between AEE:PAQ and AEE:DLW was no longer significant ($P = .91$). The relationship between MVPA:PAQ and MVPA:A was also positive ($b_1 = 0.91$; minutes per day) and significant ($P < .0001$) with a residual variance of 507.7. Although adjusting for fat mass explained a significant amount of the variance ($P = .04$), the relationship (following adjustment for fat mass) between MVPA:PAQ and MVPA:A remained significant ($P < .0001$). Sex did not explain a significant amount of the variance in any of the analyses.

The correlations between biomarkers and the PAQ, DLW, and accelerometers are presented in Table 2. Table 2 demonstrates that

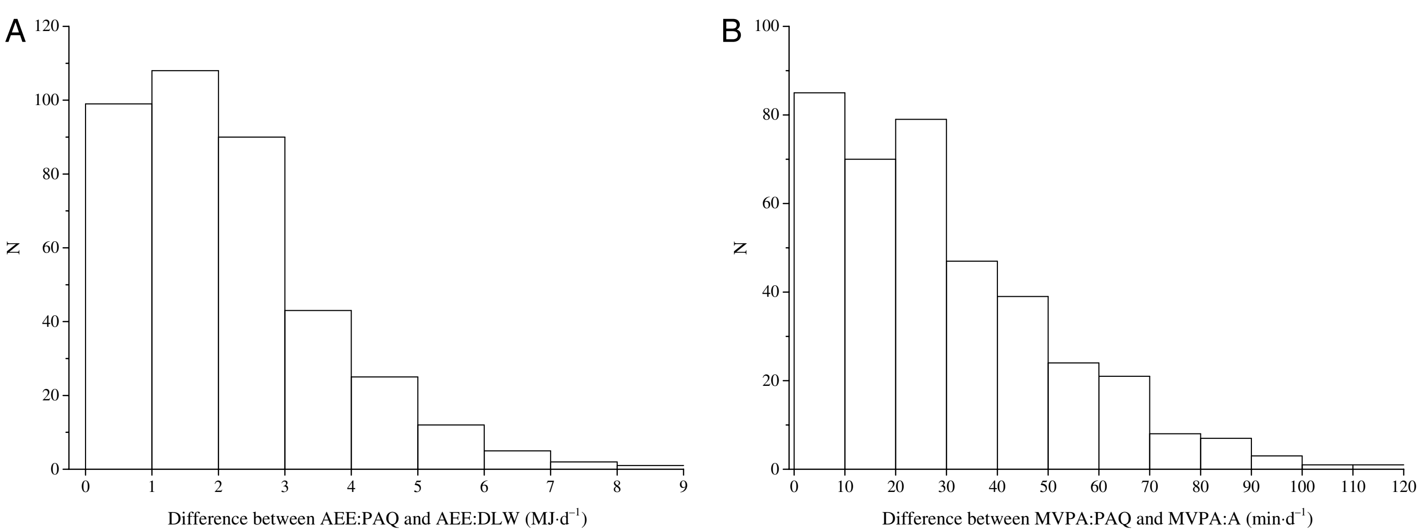

Figure 1 — Histograms of absolute differences between the PAQ and DLW for estimates of AEE (A), and PAQ and accelerometer for estimates of MVPA (B). PAQ indicates physical activity questionnaire; DLW, doubly labeled water; AEE, activity energy expenditure; MVPA, moderate to vigorous physical activity; AEE:DLW, AEE via DLW; AEE:PAQ, AEE via PAQ; MVPA:A, MVPA via accelerometer; MVPA:PAQ, MVPA via PAQ.

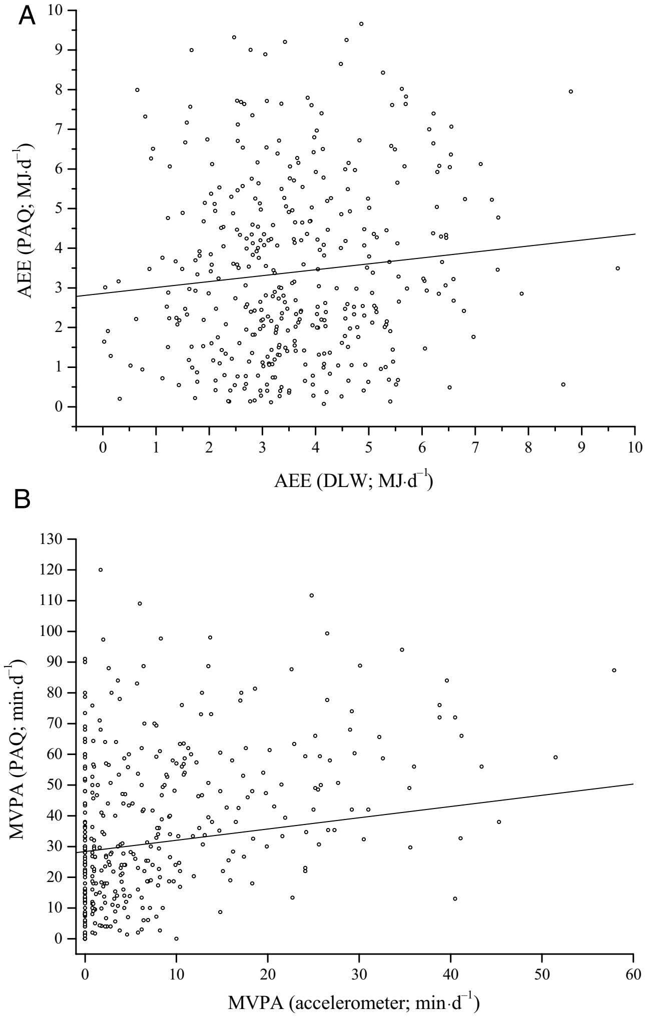

Figure 2 — Differences between the PAQ and DLW for estimates of AEE (A), and PAQ and accelerometer for estimates of MVPA (B). PAQ indicates physical activity questionnaire; DLW, doubly labeled water; AEE, activity energy expenditure; MVPA, moderate to vigorous physical activity.

10 of the 12 correlations estimated from the PAQ were lower than DLW or accelerometer. In 3 cases (AEE and body mass index, AEE and high-density lipoprotein, and MVPA and fat-free mass), the correlations were attenuated sufficiently to result in a change in the detection of statistical significance between the variables. Figure 3 plots the relationships between how participants were ranked according to AEE and MVPA. The correlation coefficient between ranking of AEE:PAQ and AEE:DLW was .09 ($P = .07$), whereas the correlation coefficient between MVPA:PAQ and MVPA:A was .40 ($P < .0001$).

Data from the PAQ indicated that 273 participants met the PAGA, and 112 participants did not meet the PAGA. However, the accelerometers (the criterion measure) estimated that only 48 participants met the PAGA, and 337 participants did not meet the recommendations.

Discussion

The results of this investigation suggest that group-level (mean) estimates of AEE from the PAQ were similar to DLW, yet MVPA was overestimated by the PAQ when compared with accelerometers. However, within-participant variability in PAQ-estimated AEE and MVPA were substantial when compared with these criterion techniques. These large measurement errors resulted in attenuated (underestimated) relationships between AEE or MVPA and certain biomarkers, essentially because the PAQ values are subject to excessive variance. The measurement errors in the estimation of MVPA resulted in the PAQ misclassifying participants who met or who did not meet the PAGA recommendations for physical activity.

The tendency for PAQs to result in biased estimates of energy expenditure and physical activity has been reported extensively elsewhere.[9,14,18,20,35] In the present study, the group-level bias for AEE was −0.2 MJ per day (5.0%) and 28.2 minutes per day (353.0%) for MVPA, with average within-participant absolute differences of 2.1 MJ per day (coefficient of variation = 71.4%) and 28.2 minutes per day (coefficient of variation = 76.7%) for AEE and MVPA, respectively. Although Figure 1 demonstrates that approximately 47% of the within-participant errors for AEE

Table 2 Comparing the Relationships Between the PAQ and DLW, and the PAQ and Accelerometers for Clinically Relevant Biomarkers

	Body fat, %	BMI, kg·m^{-2}	Fat mass, kg	Fat-free mass, kg	HDL, mg·dL^{-1}	TG, mg·dL^{-1}
Activity energy expenditure						
DLW						
r	−.14	.13	.07	.34	−.17	.08
P	.01	.01	.18	<.0001	.0001	.14
PAQ						
r	−.19	.05	−.06	.23	−.10	.08
P	.0001	.35*	.23	<.0001	.06*	.12
Moderate to vigorous physical activity						
Accelerometer						
r	−.26	−.16	−.23	.09	−.05	−.01
P	<.0001	.002	<.0001	.01	.36	.88
PAQ						
r	−.20	−.11	−.18	.06	−.02	.01
P	<.0001	.03	.001	.21*	.68	.92

Abbreviations: BMI, body mass index; DLW, doubly labeled water; HDL, high-density lipoprotein; PAQ, physical activity questionnaire; TG, triglyceride.
*Change in statistical significance between PAQ or DLW and PAQ or accelerometer.

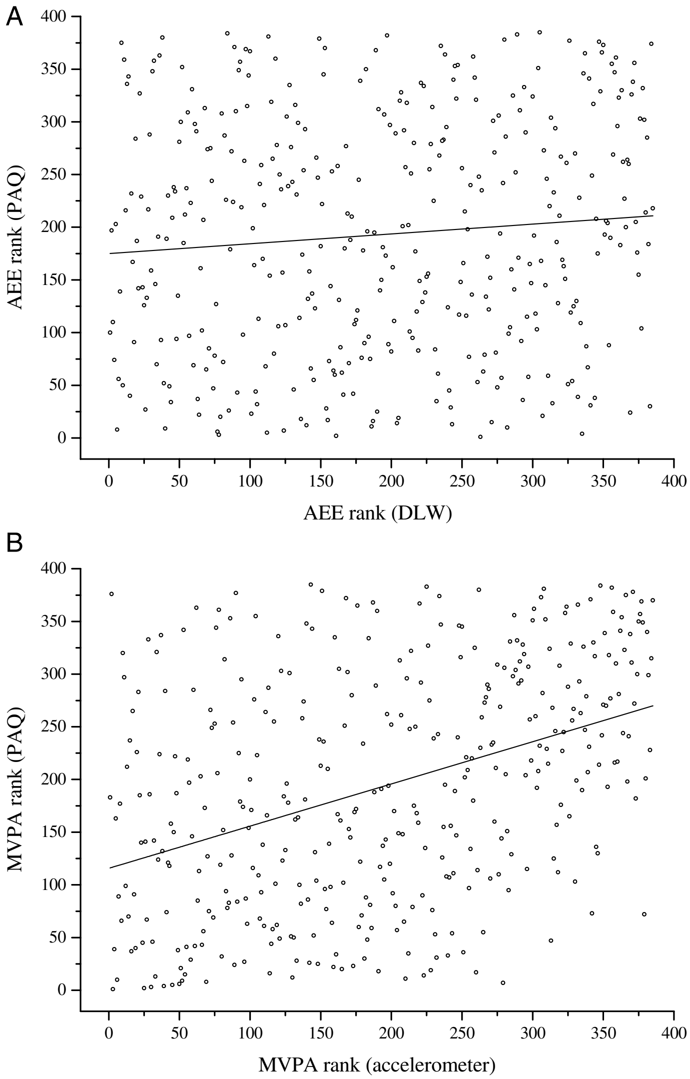

Figure 3 — Ranking of participants from the PAQ and DLW for estimates of AEE (A), and PAQ and accelerometer for estimates of MVPA (B). PAQ indicates physical activity questionnaire; DLW, doubly labeled water; AEE, activity energy expenditure; MVPA, moderate to vigorous physical activity.

were greater than 2 MJ per day, approximately 44% of the within-participant errors for MVPA were >20 minutes per day. Comparable results have been detected for the NHANES PAQ, whereby Tucker et al[17] found that reported MVPA was 398.1 minutes per week when compared with 63.7 minutes per week via accelerometer. (This difference was 197.4 minutes per week in the present study.) Our results are also consistent with Prince et al,[14] who suggested that self-reported physical activity is more likely to be underreported when compared with DLW and overreported when compared with accelerometers.

Although energy expenditure and physical activity data are often used in regression models to predict relationships with a variety of biomarkers for disease, the presence of measurement error likely results in attenuation of the true relationships between these variables.[36,37] Table 2 suggests that although positive and statistically significant relationships between AEE:PAQ and AEE:DLW (not significant after taking fat-free mass into account), and MVPA:PAQ and MVPA:A can be observed, utilizing AEE and MVPA from the PAQ resulted in attenuated relationships with 10 of the 12 biomarkers measured in this study. In 3 of these comparisons, statistically significant relationships detected with the criterion measures were not detected by the PAQ. The potential for the NHANES PAQ to underestimate relationships with disease has also been reported by Tooze et al.[36] Despite the underestimate of biomarkers with the PAQ when compared with DLW and accelerometers, it is worth noting that the amount of explained variance in all of the comparisons was small (Table 2).

The presence of measurement error in this study provides additional evidence that the true relationships between energy expenditure, physical activity, and disease may be underestimated with PAQs.[1] Matthews et al[7] recognized the importance of attenuation but suggested that PAQs may still provide useful information because of their ability to "rank" sedentary or active individuals. In other words, it is possible that imprecise estimates of AEE and MVPA from PAQs may still useful in predicting disease as long as participants are "ranked" against each other accurately. However, our results demonstrated that the rankings of participants correlated poorly between the PAQ and the criterion measures.

Although this investigation demonstrates concerns about the ability of the PAQ to predict AEE and MVPA, it is possible that these data may alternatively be used to categorize individuals who meet national physical activity recommendations. However, only 12.5% participants met the PAGA according to the accelerometer (the criterion estimate), compared with 70.9% via PAQ. The probability of the NHANES PAQ to misclassify individuals for meeting national physical activity guidelines when compared with accelerometers has also been reported in 2 previous reports.[6,17] Tucker et al[17] reported that 59.6% of participants met the PAGA by the NHANES PAQ, compared with only 8.2% from accelerometers.

There are a number of strengths and limitations of the present study. For strengths, this study included a relatively large sample size and simultaneously utilized DLW and accelerometers to estimate both AEE and MVPA (also comparing them in the same units), and utilized measured REE via indirect calorimetry.[9,14,25] We took these results further by characterizing the measurement error from the PAQ. However, it is not well understood how NHANES calculates TEE or REE, or screens out implausible values. (In total, 101 of 486 of the participants were excluded due to implausible PAQ results in this study.) The impacts of these exclusions on the analyses are difficult to determine and may influence the generalizability of the results; however, the differences in demographics between those excluded and included in the analyses were reasonably small. Finally, this study operates under the assumption that DLW and accelerometers are "gold standards" of within-participant AEE and MVPA (respectively) and that any disagreements between these estimates and the PAQ are due solely to the PAQ.[9,26] However, this is not likely the case, particularly when utilizing DLW as a measure of AEE.[12]

In conclusion, the NHANES PAQ is subject to a considerable amount of bias and variance when used to estimate AEE and MVPA. The implications of these results take many forms, such as impacts on the development of health policy and understanding the progress of physical activity promotion efforts over time. For example, the relationships between physical activity and disease have often been used to guide physical activity policy development, yet these results suggest that the application of PAQ data may result in underestimates of the true relationships between AEE or MVPA and disease. Also, the tendency for the NHANES PAQ to overestimate MVPA results in misclassification of individuals who meet the PAGA. Future studies should be designed to refine this approach by the implementation of calibration studies[27] and measurement error correction models that may improve the accuracy of PAQs, such as investigating the usefulness of covariates. Finally, a more complete understanding of the reasons why participants report implausible results on the PAQ could increase the precision of AEE and MVPA estimates.

Acknowledgment

This work was supported by the US Department of Agriculture-Agricultural Research Service.

References

1. Manini TM, Everhart JE, Patel KV, et al. Daily activity energy expenditure and mortality among older adults. *JAMA*. 2006;296(2):171–179. PubMed doi:10.1001/jama.296.2.171
2. Hu FB, Willett WC, Li T, Stampfer MJ, Colditz GA, Manson JE. Adiposity as compared with physical activity in predicting mortality among women. *N Engl J Med*. 2004;351(26):2694–2703. PubMed doi:10.1056/NEJMoa042135
3. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004;291(10):1238–1245. PubMed doi:10.1001/jama.291.10.1238
4. Jakicic JM, Otto AD. Physical activity considerations for the treatment and prevention of obesity. *Am J Clin Nutr*. 2005;82(1)(suppl):226S–229S. PubMed.
5. US Department of Health and Human Services. *Physical Activity Guidelines for Americans*. Washington, DC: US Department of Health and Human Services; 2008:vi–vii.
6. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181–188. PubMed doi:10.1249/mss.0b013e31815a51b3
7. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving self-reports of active and sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev*. 2012;40(3):118–126. PubMed.
8. Ainsworth BE, Caspersen CJ, Matthews CE, Masse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. *J Phys Act Health*. 2012;9-(suppl1):S76–S84. doi:10.1123/jpah.9.s1.s76
9. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr*. 2008;87(2):279–291. PubMed.
10. Esliger DW, Tremblay MS. Physical activity and inactivity profiling: the next generation. *Can J Public Health*. 2007;98(suppl 2):S195–S207.
11. Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr*. 1988;118(11):1278–1289. PubMed doi:10.1093/jn/118.11.1278
12. Speakman JR. The history and theory of the doubly labeled water technique. *Am J Clin Nutr*. 1998;68(4):932S–938S. PubMed doi:10.1093/ajcn/68.4.932S
13. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes Rev*. 2013;14(6):451–462. PubMed doi:10.1111/obr.12021
14. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act*. 2008;5:56. PubMed doi:10.1186/1479-5868-5-56
15. Paul DR, Kramer M, Stote KS, Baer DJ. Determinants of variance in the habitual physical activity of overweight adults. *J Phys Act Health*. 2015;12(5):680–685. PubMed doi:10.1123/jpah.2013-0269
16. Melanson EL Jr, Freedson PS. Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr*. 1996;36(5):385–396. PubMed doi:10.1080/10408399609527732
17. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. *Am J Prev Med*. 2011;40(4):454–461. PubMed doi:10.1016/j.amepre.2010.12.016
18. Csizmadi I, Neilson HK, Kopciuk KA, et al. The Sedentary Time and Activity Reporting Questionnaire (STAR-Q): reliability and validity against doubly labeled water and 7-day activity diaries. *Am J Epidemiol*. 2014;180(4):424–435. PubMed doi:10.1093/aje/kwu150
19. Neuhouser ML, Di C, Tinker LF, et al. Physical activity assessment: biomarkers and self-report of activity-related energy expenditure in the WHI. *Am J Epidemiol*. 2013;177(6):576–585. PubMed doi:10.1093/aje/kws269
20. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. *Am J Clin Nutr*. 2010;91(1):106–114. PubMed doi:10.3945/ajcn.2009.28432
21. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *J Phys Act Health*. 2009;6(6):790–804. PubMed doi:10.1123/jpah.6.6.790
22. Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc*. 2001;33(7):1233–1240. PubMed doi:10.1097/00005768-200107000-00024
23. De Vera MA, Ratzlaff C, Doerfling P, Kopec J. Reliability and validity of an internet-based questionnaire measuring lifetime physical activity. *Am J Epidemiol*. 2010;172(10):1190–1198. PubMed doi:10.1093/aje/kwq273
24. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381–1395. PubMed doi:10.1249/01.MSS.0000078924.61453.FB
25. van Poppel MN, Chinapaw MJ, Mokkink LB, van Mechelen W, Terwee CB. Physical activity questionnaires for adults: a systematic review of measurement properties. *Sports Med*. 2010;40(7):565–600. PubMed doi:10.2165/11531930-000000000-00000
26. Winkler E, Waters L, Eakin E, Fjeldsoe B, Owen N, Reeves M. Is measurement error altered by participation in a physical activity intervention? *Med Sci Sports Exerc*. 2013;45(5):1004–1011. PubMed doi:10.1249/MSS.0b013e31827ccf7d
27. Zheng C, Beresford SA, Van Horn L, et al. Simultaneous association of total energy consumption and activity-related energy expenditure with risks of cardiovascular disease, cancer, and diabetes among postmenopausal women. *Am J Epidemiol*. 2014;180(5):526–535. PubMed doi:10.1093/aje/kwu152
28. Westerterp KR. Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol*. 2009;105(6):823–828. PubMed doi:10.1007/s00421-009-1000-2
29. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr*. 2008;88(2):324–332. PubMed

30. Schoeller DA, Tylavsky FA, Baer DJ, et al. QDR 4500A dual-energy X-ray absorptiometer underestimates fat mass in comparison with criterion methods in adults. *Am J Clin Nutr*. 2005;81(5):1018–1025. PubMed
31. Paul DR, Novotny JA, Rumpler WV. Effects of the interaction of sex and food intake on the relation between energy expenditure and body composition. *Am J Clin Nutr*. 2004;79(3):385–389. PubMed doi:10.1093/ajcn/79.3.385
32. Food and Nutrition Board. *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids*. Washington, DC: National Academies Press; 2005:157–158.
33. Paul DR, Kramer M, Stote KS, et al. Estimates of adherence and error analysis of physical activity data collected via accelerometry in a large study of free-living adults. *BMC Med Res Methodol*. 2008;8:38. PubMed doi:10.1186/1471-2288-8-38
34. Gordon B, Chen S, Durstine JL. The effects of exercise training on the traditional lipid profile and beyond. *Curr Sports Med Rep*. 2014;13(4):253–259. PubMed doi:10.1249/JSR.0000000000000073
35. Loney T, Standage M, Thompson D, Sebire SJ, Cumming S. Self-report vs. objectively assessed physical activity: which is right for public health? *J Phys Act Health*. 2011;8(1):62–70. PubMed doi:10.1123/jpah.8.1.62
36. Tooze JA, Troiano RP, Carroll RJ, Moshfegh AJ, Freedman LS. A measurement error model for physical activity level as measured by a questionnaire with application to the 1999-2006 NHANES questionnaire. *Am J Epidemiol*. 2013;177(11):1199–1208. PubMed doi:10.1093/aje/kws379
37. Schatzkin A, Subar AF, Moore S, et al. Observational epidemiologic studies of nutrition and cancer: the next generation (with better observation). *Cancer Epidemiol Biomarkers Prev*. 2009;18(4):1026–1032. PubMed doi:10.1158/1055-9965.EPI-08-1129